

Claims

We claim:

- 1 1. A method for temporally processing an input video including a plurality
2 of input frames, each of the input frames having an associated input frame
3 play time, and the input video having a total input video play time that is a
4 sum of the input frame play times of all of the input frames, comprising:
5 classifying each of the plurality of input frames according to a content
6 characteristic of each of the input frames; and
7 allocating an output frame play time to each of the plurality of input
8 frames that is based on the classified content characteristic of each of the
9 input frames to generate a plurality of output frames.
- 1 2. The method of claim 1 wherein the content characteristic is based on low-
2 level features of each of the input frames.
- 1 3. The method of claim 1 wherein the low-level features are selected from a
2 group consisting of motion vectors, color, texture, brightness, contrast,
3 spectral parameters, local and global motion, activity, trajectory, speed,
4 acceleration, object shape, object size, number of objects, number of faces,
5 pitch, volume, tempo, and combinations thereof.
- 1 4. The method of claim 1 wherein the content characteristic is based on
2 high-level features of each of the input frames.

1 5. The method of claim 1 wherein the high-level features are selected from a
2 group consisting of genre, dramatic intensity, humor content, action level,
3 beauty, lyricism, musical intensity, educational quality, profundity, nudity,
4 linguistic class, and combinations thereof.

1 6. The method of claim 1 wherein the allocating of the play time is
2 dynamically varied while processing the video.

1 7. The method of claim 1 wherein the allocated output frame play time of
2 each of the output frames is determined by sampling the input frames.

1 8. The method of claim 7 wherein the sampling is a down-sampling of the
2 input frames.

1 9. The method of claim 7 wherein the sampling is an up-sampling of the
2 input frames.

1 10. The method of claim 9 wherein up-sampled output frames are
2 interpolated from the input frames.

1 11. The method of claim 7 wherein the sampling is a combination of down-
2 sampling and up-sampling of the input frames.

1 12. The method of claim 1 wherein the allocated output frame play time of
2 each of the output frames is determined by an output frame rate of the output
3 frame.

1 13. The method of claim 12 wherein the output frame rate is increased for
2 selected input frames.

1 14. The method of claim 12 wherein the output frame rate is decreased for
2 selected input frames.

1 15. The method of claim 1 further comprising:
2 measuring the content characteristics of each of the plurality of input
3 frames to determine the classification.

1 16. The method of claim 15 further comprising:
2 computing a statistical moment for the measured characteristics to
3 determine the classification.

1 17. The method of claim 1 wherein the allocation of play time is based on a
2 constant level of motion activity in the output video.

1 18. The method of claim 1 wherein the allocation of play time is based on a
2 guaranteed minimum level of activity in the output video.

1 19. The method of claim 1 further comprising:
2 partitioning the input video into a plurality of segments, and
3 processing the input video on a per segment basis.

1 20. The method of claim 1 wherein still frames are selected for the output
2 video when the allocated output frame play time exceeds a temporal Nyquist
3 limit.

1 21. The method of claim 1 further comprising:
2 allocating a total output video play time; and
3 allocating the output frame play times so that a sum of the output
4 frame play times of the plurality of output frames is equal to the total output
5 video play time of the output video.

1 22. The method of claim 1 wherein the allocated play time of a particular
2 frame can range on a continuum from zero time to a length of time of the
3 output video.

1 23. The method of claim 1 wherein the allocation of play time is based on a
2 motion activity in the output video, and a measure of motion activity is an
3 average of magnitudes of motion vectors of the frames.

1 24. The method of claim 23 where the average motion vector magnitude \hat{r} of
2 the input video of N frames is expressed as:

3
$$\hat{r} = \left(\frac{1}{N}\right) \sum_{i=1}^N r_i ,$$

4 where an average motion vector magnitude of frame i is r_i .

1 25. The method of claim 24 wherein a relationship between a length L_{output} of
2 the output video and a length L_{input} of the input video is expressed as

$$L_{output} = \frac{\hat{r}}{r_{target}} L_{input}$$

for a target level of motion activity r_{target} in the output video.

26. The method of claim 25 further comprising:

classifying all of frames j of the input video having the motion activity equal to or higher than a targeted level of minimum motion activity into a first set S_{higher} having a length L_{higher} ;

classifying all of frames k of the input video having the motion activity lower than the targeted level of minimum motion activity into a second S_{lower} having a length L_{lower} ;

summing $L_{higher} + L_{lower}$ to determine a L_{input} of the input video to determine a length of the output video by

$$L_{output} = \left(\frac{\hat{r}_{lower}}{r_{target}} \right) L_{lower} + L_{higher} .$$

27. A system for temporally processing an input video including a plurality of input frames, each of the input frames having an associated input frame play time, and the input video having a total input video play time that is a sum of the input frame play times of all of the input frames, comprising:

means for classifying each of the plurality of input frames according to a content characteristic of each of the input frames;

means for allocating a total output video play time; and

means for allocating an output frame play time to each of the plurality of input frames that is based on the classified content characteristic of each of the input frames to generate a plurality of output frames so that a sum of

